DOCUMENT RESUME

ED 043 564 SP 004 144

AUTHOR Ornstein, Allan C.
TITLE Research on Teacher Behavior, Approaches, Limitations, and Recommendations.
INSTITUTION New York Univ., N.Y.
PUB DATE 70
NOTE 54p.

EDRS PRICE EDRS Price MF-$0.25 HC-$2.80
DESCRIPTORS *Educational Research, Measurement Techniques, Research Criteria, *Research Problems, *Teacher Behavior

ABSTRACT

This review of research on teacher behavior covers 200 references cited in the bibliography, the bulk of them published in the 1960's. A short introductory section on the status of research on teacher behavior notes that it is "voluminous and contradictory." The main body of the paper is organized in terms of the reason why research on teacher behavior yields no clear conclusions, a threefold problem involving the theoretical framework whereby an investigator approaches teacher behavior research: 1) systematizing research on teacher behavior--into three categories: models systems, instructional processes, and teacher behavior characteristics; 2) identifying teacher behaviors; and 3) measuring teacher behaviors--by observation (three types), student behavior and achievement, tests based on recall, and psychological tests. Another section details some additional problems related to research on teacher behavior, e.g., its noncumulative nature (researchers measuring different phenomena and variables, using different terms, methods, and assumptions), questions as to whether or not teacher behavior research may be beyond scientific analysis, and problems regarding teacher-researcher relations and practical application of theory. The conclusion section lists 27 recommendations for future research on teacher behavior. (JS)

ED 043564

RESEARCH ON TEACHER BEHAVIOR: APPROACHES, LIMITATIONS, AND RECOMMENDATIONS*

Allan C. Ornstein

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

The Status of Research on Teacher Behavior

The research on teacher behavior is voluminous and contradictory. Biddle (1964), Eisner (1963), Flanders (1964), and D. W. Johnson (1969) maintain that the problem is so complex that no one knows or agrees upon what a competent teacher is. Broudy (1969) contends that we "can define good teaching any way we like [p. 583]." Morris (1969) affirms that since we are unable to define a good teacher, it is injudicious to formulate scientifically based generalizations about good teacher behavior. Eisner (1963) and Macdonald and Zaret (1968) believe that we lack the ability to evaluate the symbolic aspects of teaching. Cronbach (1966) and P. W. Jackson (1968) maintain that we lack sufficient knowledge about learning to evaluate teacher behavior or instruction adequately. Goheen (1966) points out that teacher behavior cannot be defined and analyzed, and therefore, "there will always be teachers who will break all the rules and yet be profoundly successful [p. 221]." Kerlinger (1967) asserts that no single teacher can possibly possess all the traits listed in several studies. Eisner (1963) and J. P. Jackson (1968) contend that teachers are relatively unaffected by teacher-behavior research. Flanders (1960) and French (1961) affirm that research findings are not applicable to a specific classroom situation - teachers or students. Gage (1968)

---

*This paper is primarily based on the author's first draft of Chapter 3, "Conceptual Framework," in Selected Teacher Behavior Attributes Rated as Desirable by Ninth-Grade Disadvantaged Students and Ninth-Grade Teachers of the Disadvantaged. Unpublished doctoral dissertation. New York University, expected October, 1970.

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Allan C. Ornstein
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

SP004144

alleges that teacher-behavior research does not make good sense or "hang together" in a comprehensible way. Guba and Getzels (1955) aver that teacher-behavior research has failed "to produce generally meaningful and useful results [p. 330]." Wall (1969) declares that the "findings ... are either confirmations of 'common sense' or manifestly absurd [p. 163]." Biddle and Ellena (1964) state that acceptable findings are often later refudiated. P. W. Jackson (1966) believes that the few discoveries up to now are "pitifully small" in proportion to the outlay in time and effort. Tanner (1969) is of the opinion that "emphasis on teacher behavior has gone to excess [p. 366]," and Biddle (1964) claims it is "becoming unmanageable [p. 2]." Elsewhere, Biddle (1967) affirms that "the investigators[1] themselves do not know what to make of [their] findings [p. 348]." Berelson and Steiner (1964) summed and dismissed research on teacher behavior in five words - "there are no clear conclusions [p. 441]." Why? The problem is threefold: involving the "theoretical framework," whereby the invesitigator (1) systematizes the concepts of the research, as well as the "criterion," a twofold problem, whereby the investigator (2) identifies and (3) measures teacher behavior.

Systematizing Research on Teacher Behavior

Methods for organizing teacher-behavior research generally fall into one of three categories: models systems, instructional processes, and teacher-behavior characteristics. Of the three, the model-systems approach is the most sophisticated, and it may include instructional processes and/or teacher-behavior characteristics, along with other inputs or outputs. The instructional processes approach, next in sophistication, views teaching as a continuous interaction process between teacher and students, and evaluates this interaction, namely teacher and student behavior, by observation. The

---

[1] The terms "investigators" and "researchers" are used interchangeably by the author in this paper.

teacher-behavior-characteristics approach is the least sophisticated, either constituting a study in itself or comprising, in part, one of the two more sophisticated approaches.

Model systems. According to Ryans (1963b), the model system is an indentifiable scheme of complex but organized and interrelated elements and/or subelements which function as a coordinated prototype. The model directs attention to the systematic nature of teaching and learning. A few examples of the model systems are: (A) Flanders (1960), involving the teacher's (1) authority, (2) goals, (3) interaction, and (4) flexibility. (B) Jensen (1955), comprising seven categories of class productivity and class cohesiveness: (1) problem-solving, (2) authority-leadership, (3) power, (4) friendship, (5) personal prestige, (6) sex, and (7) privilege. (C) Getzels and Thelen (1960) who view the classroom as a social system - governed by institutional roles and individual needs and three teaching styles, namely, (1) nomothetic, (2) idiographic, and (3) transactional - which indicate to what extent the teacher gravitates toward the institution or the individual. (D) Ryans' (1963b) communication of information model, constituting (1) classifying, (2) evaluating, (3) decision-making, (4) ordering, and (5) transmitting. And (E) Biddle (1964), embracing cause-and-effect factors for teacher effectiveness - (1) formative experiences, (2) teacher properties, (3) teacher behaviors, (4) immediate effects, (5) long-term consequences, (6) classroom situations, and (7) school and community contexts.

The model system tends to be all-embracing - attempting to include every variable - and according to Gage and Unruh (1967), making it too unwieldy for effective research. If it does include every variable, or nearly every variable, which is still difficult to imagine, it follows, then, the model,

without the investigator's knowledge or ability to make subtle distinctions, may often manipulate insignificant variables and/or inconsequential events, moreover, assign them equal weights. Even if the investigator discerns which variables are more important, the validity can be questioned, and even assuming different weights are assigned, they too can be questioned.[2]

Gage and Unruh (1967) and Siegel and Siegel (1967) assert that the number of combinations make the model approach unmanageable; likewise, Cronbach (1967) and Stolurow (1965) affirm that teacher-student interactions are too many and too complex to simplify into a model. Atkin (1967-68) contends that the model system erroneously reinforces a fragmentized view of teaching and such a view transmutes the model into trivia.[3] P. W. Jackson (1966, 1968) maintains that teacher behavior cannot be processed into a model; it is too spontaneous and uncontollable.[4] Atkin (1967-68) declares that there are too many subtleties inherent in teaching, making the model illogical - though its appeal is based on logic. Also, Atkin affirms that the model approach is based on precise calcuations of inputs and specifications of performance; however, teacher-behavior inputs and specifications are obscure and vary with different researchers - producing, according to this author, a distorted, deficient, and/or inconceivable model. Gage and Unruh (1967) aver the model system to be merely "metaphorical, not to be taken literally [p. 361]." In this connection, none of the five model specimens seem to utilize empirical data or have been put to the test or consumated in a classroom situation.

---

2 *Infra*, p. 25, fn. # 16.

3 See pp. 38-39. However, recommendation # 14 suggests that teacher behavior components be broken down so that the variables will be more manageable. Reccomendation # 15 suggests a micro-analytical approach for measuring teacher behavior. Both recommendations might be considered as fragmentizing teacher-behavior research.

4 *Infra*, p. 30.

Instructional processes. The instructional process may be divided into two groups. One group tends to be descriptive and based on nonsystematic observations. The observer usually enters the classroom, takes notes, develops insights, and is at liberty to analyze nearly any facet of the teaching-learning process; no empirical data is forthcoming. Exemplifying this type of teacher-behavior analysis are Eddy (1967), Henry (1955), Holt (1964), and P. W. Jackson (1968). These observers tend to be indicative of the anti-teacher syndrome; one might contend that have preconceived, negative notions about the teacher and what goes on in the classroom, and since this approach does not call for reliability or validity, they seem merely to fit their observations into their biases. Since this method ignores research methodology, it is not elaborated upon in this paper - merely noted so as to distinguish between the second type of instructional process. This second approach tends to be analytical and based on systematic observation in which the observer(s) tests preconceived hypotheses, teacher-student interaction, and/or teacher-behavior characteristics.

The second type of instructional process, according to Biddle (1967), is usually decribed and evaluated by a defined, but abstract, unit of measurement, varying in size and involving a sequence of responses, or according to this author, some kind of distinct or prescribed behavior by the teacher and student(s): "moves" (Bellack, 1966; Meux & Smith, 1964), "acts" (Flanders, 1965; Macdonald & Zaret, 1968), or "messages" (Galloway, 1962). A series of responses or behaviors usually constitutes a separate set: "episode" (Meux & Smith, 1964; B. O. Smith, 1964), "cycle" (Bellack, 1966), "pattern" (Flanders, 1965), "incident" (Macdonald & Zaret, 1968), or "communication" (Galloway, 1962).

A few examples of this instructional process suffice; they are: (A) Meux

and Smith (1964) and B. O. Smith (1964), who devised 13 categories for placing teacher-student "moves" into "episodes" or "monologues;" (B) Bellack (1966), who classified a set of four "moves" by teachers and students - (1) structuring, (2) soliciting, (3) responding, and (4) reacting - into a "cycle;" (C) Flanders (1965), who devised a nine-point interaction scale involving statements or "acts" by teachers and students, thereby classifying "indirect" and "direct" teacher behavior or "patterns;" (D) Macdonald and Zaret (1968), who classified varbal "acts" into "opening" teacher-student behavior or "incidents," leading to productive learning, and "closing" teacher-student behavior, leading to reproductive learning. And (E) Galloway (1962), who classified three kinds of "messages" - (1) facial expressions, (2) actions, and (3) vocal language - into seven possible categories of nonverbal "communication" toward students.

Since the instructional process is usually dependent, at least in part, on categorizing teacher-behavior characteristics and assessing teacher and student behavior by observations, the discussion below of the limitations of the aforementioned "criteria" are applicable, too, to the instructional process. Thus, part of the discussion that is to follow, specifically "identifying teacher behaviors," "observations," and "student behavior," is relevant to the instructional process.

Identifying Teacher Behaviors

Among the reams of research on teacher behavior, there are many options for choosing teacher-behavior characteristics (Flanders, 1965; Ryans, 1963a);[5] this in itself, causes a problem, that is, Barr, Eustice, and Noe (1955), Gage (1968), Ryans (1960, 1964), B. O. Smith (1967), and Start (1966) contend

---

[5] The list appears endless, therefore, it would be fruitless to list examples, as with the discussion on model systems and instructional processes.

that our inability to define or agree on which teacher behaviors constitute "good" teacher behavior or "effective" teaching, has confused researchers and/or caused inconsistencies among the research findings. Flanders (1964) affirms that a particular pattern of teacher behavior cannot be advocated, and Broudy (1969) and Eisner (1963) contend that these behavior patterns cannot be reduced to a formula or rule. Also, Berelson and Steiner (1964), Biddle (1964) and Rosencranz and Biddle (1964) point out that there is no consistent relationship between teacher behavior and teaching.

Some investigators, for example, Bettelheim (1961), Hargadon (1966), Rogers (1959), Sheviakov and Redl (1956), and Stavsky (1957) contend - either directly or indirectly - that it is fruitless to try to identify "good" teacher behavior, because teaching involves an interpersonal relationship - human behavior - between teacher and student(s) - which must be described and analyzed. Nevertheless, these invesitgators fail to provide an empirical method for evaluating their recommendations or for conducting research.

According to the American Educational Research Committee[6] (1952) and Biddle (1964) confusion over a variety of terms, such as "teacher behavior," "teacher-behavior characteristics," "teacher traits," "teacher personality," "teacher competence," "teacher performance," etc. add to the general problem. Even worse, according to this author, the definition and usuage of these terms vary among different researchers.

AERC (1952), Biddle and Ellena (1964), Gage (1968), Ryans (1964), and Wehling and Charters (1969) maintain that there are too many teacher behaviors

---

[6] Hereinafter cited as AERC.

to analyze or measure. AERC (1952), Getzels and Jackson (1963), Rosencranz and Biddle (1964), and Ryans (1964) are of the opinion that there is lacking of agreement upon a method for evaluating teacher behavior. Biddle (1964, 1967), Foa (1965), Gage and Unruh (1967), and Hyman (1968) affirm there are no clear or acceptable methods for categorizing and/or identifying teacher behaviors. Carroll (1964), Klein (1969), and Withall (1951) assert that teacher-behavior categories are vague and ill-defined. Ryans (1963a) believes that teacher behaviors are not generalizable to other teachers. Hyman (1968) and Meux and Smith (1964) alledge that there is difficulty in classifying teacher behaviors into proper [and valid] dimensions; teacher behavior from one study often cannot be categorized into the same dimension in another study. On the other hand, Meux and Smith (1964), Ryans (1960), and Wehling and Charters (1969) affirm that different teacher behaviors categorized into a specific dimension, despite their "independence," are often related either logically or statistically. Carroll (1964), Flanders (1964), Hyman (1968), B.O. Smith (1967), and Tanner (1969) believe that the validity or "independence" of teacher behaviors which are categorized into dimensions are likely to decrease with the increase of dimensions - overlapping increases, while mutual exclusiveness decreases. Yet most of the aforementioned investigators point out that if the teacher-behavior dimensions are decreased, the findings are oversimplified and little worthwhile data are forthcoming. Thus, Flanders (1964) and Klein (1969) question whether a set of criteria can be developed to provide sufficient properties for classifying teacher behaviors.

Biddle (1967) and Perkins (1964) maintain that there are too many "similarities" and "dissimilarities" among the different teacher-behavior categories, causing serious and confusing analytical problems. For example, a teacher

that "gives direction" would be exhibiting "direct behavior" by Flanders (1965), "controlling behavior" by Hughes (1962, 1965) and Ornstein (1970), "routine" behavior by Gallaher and Aschner (1963,1965), "directing and managing" behavior" by Meux and Smith (1964), and "responsible" behavior by Ryans (1960). These different teacher-behavior categories, although somewhat similar, tend to invalidate comparisons between different studies. Meux and Smith (1964) allege that a particular teacher behavior judged to be "effective" in one study can be judged "ineffective" in another study. According to Getzels and Jackson (1963), the only consistencies are the obvious teacher behaviors; for example, "friendly" behavior is indicative of a "good" teacher and the opposite-type behavior is indicative of a "poor" teacher.

Biddle (1964) points out that there are thousands of descriptive words that may be applied for describing or classifying teacher behavior. For example, with one teacher behavior alone, namely, verbal behavior, Flanders (1965) employed 7 different examples while Zohorok (1968) used 175 different examples. Assuming content validity in both cases, who is right, and who determines who is right?, judges are biased, so is this reader. Into how many different components can verbal behavior, for that matter any type of teacher behavior, be subdivided? No one really seems to know, at least agree! Similarly, Meux and Smith (1964), Ryans (1964), and J. J. Turner (1964) are of the opinion that linguistic usuage, confusion over words, and/or interchangeability of words cause difficulties concerning agreement on operational or behavioral meanings of teacher-behavior categories, or, according to Jenkins (1960) and Perkins (1964), in the way in which teacher behavior

---

*From the author's unpublished doctoral dissertation. See footnote on p. 1.

occurrs, as well as the nature and scope of the behavior. For example, this author* used "welcomes and is respectful of views other than own" as a behavior phrase to help describe Affective Teacher Behavior. A similar teacher behavior, "sincere sympathy with a pupil's viewpoint [p. 88]," is categorized by Ryans (1960) as Understanding Behavior. Dumas (1966) ranked "sympathy with pupil viewpoint [p. 24]" with Empathy. Medley and Mitzel (1963) identified "tried to see pupil point of view [p. 276]" with Teacher Climate. Remmers (1963), reviewing different rating scales, reported "accepted students' viewpoint with open mind [p. 342]" under Adequacy of Relations with Students. Sontag (1968) itemized "shows interest in the viewpoint of pupils [p. 395]" with Concern for Students. Jersild (1940) linked "permitted expression of opinion [p. 144]" with Teacher Performance. This type of discrepancy, this inability to agree upon operational terms, causes a lack of generalizability in the findings, as well as causes the research and related literature to be misleading. As a beginning, this seems to indicate the need for agreed upon, dictionary-type definitions of teacher behaviors.[7]

Even when there is agreement on "good" teacher behavior, it is wrong to assume that there is common meaning to the words used to describe them. Teacher-behavior concepts and definitions have different meanings with different groups or subjects - for example, students, teachers, supervisors - in part, because of their different roles (Bellack, 1966; Rosencranz & Biddle, 1964; Smith & Geoffrey, 1965); moreover, even within the same group of subjects (Cliff, 1968; McCallon & Dumas, 1967; McNeil, 1967). Also, this problem is evident with the different investigators, themselves, even though they often attempt some kind of acceptable** validity. For example, this author

---

*Ibid.

7 See p. 36, recommendation # 4a.

**From the author's unpublished doctoral dissertation. See footnote on p. 1.

used items for classifying cognitive teacher behavior. Gallacher and Aschner (1963, 1965) organized the same teacher behavior into four dimensions, based on the Guilford (1966) model of intellect, along with 11 subdimensions and 14 items to illustrate the 2 types of dimensions. Maisa (1965) organized cognitive teacher behavior into 6 dimensions, based on the Bloom et al (1956) taxonomy of educational objectives, along with 17 items to illustrate the 6 dimensions. Warren (1968) presented 40 items to evaluate the same teacher behavior. Not only do almost all the specific items differ among these investigators, but it becomes incongrous when it is pointed out that, with the exception of Gallacher and Aschner, the other investigators solely (Maisa and Warren) or largely (Ornstein) refer to Bloom for purposes of defining cognitive teacher behavior; their definitions being similar. In this connection, Biddle (1964) remarks that the investgators have their own vocabulary for defining specific teacher behaviors.

Biddle (1967), J. M. Jackson (1960), and R. G. Turner (1964) maintain that judgments about teacher behavior are socially biased. Atkin (1967-68), Glazer (1963), Gibb (1960), and Stake (1967) believe that teacher behavior varies with the nature of goals, and therefore, according to this author, comparisons of studies that fail to take this into account are misleading, and most studies fail to do it. Also, Atkin (1967-68) maintains that teacher behaviors involve values and social outcomes which cannot be quantified.

Finally, according to AERC (1952), Biddle (1964), Broudy (1969), Gage (1968), Hearn (1953), and Ryans (1964), there is no adequate criterion against which a list of teacher behaviors can be validated. According to B. O. Smith (1967), it is "inappropriate to ask whether ... a system is a true classification of the relevant phenomena." The most we can hope is

that the teacher behaviors are "wise" and "useful" and that the method for classifying them is relatively clear - bearing in mind that "this condition may never be completely satisfied by any system [p. 67]."

Measuring Teacher Behaviors

In addition to being unable to agree on a list of teacher behaviors, there is a lack of agreement, according to AERC (1952), Biddle and Ellena (1964), Gage (1968), and Getzels and Jackson (1963), on how to measure teacher behavior. Furthermore, the methods for measuring teacher behavior seem questionable in terms of reliability and/or validity (Biddle, 1964; Cronbach, 1963; Getzels & Jackson, 1963), which will be discussed below. According to Ryans (1963a), the reliability and validity of measurements of teacher behavior should be considered only relative to a defined situation, which in turn, according to this author, yields relatively ungeneralizable findings. With this, let us proceed to discuss methods for assessing and/or correlating teacher behavior; they fall into four broad areas: (1) observations, (2) student behavior and achievement, (3) tests based on recall, and (4) psychological tests.

Observations. Biddle (1967) classifies observations into three types: (1) post-session evaluation, whereby the observer makes broad evaluations of what went on after the class session is finished, (2) sign observation, whereby the observer rates a specific list of behaviors by some specific unit of time, i. e., "moves," "acts," etc., and (3) categorical observation, whereby the observer uses a scale to rate a specific list of teacher behaviors. In general, according to this author, all three techniqies are somewhat biased and deficient, which will be discussed below.

Paraskevopoulous (1968), and Veldman and Peck (1963) point out that observations of teacher behavior are limited because of the small number

upon which a rating is based. Hearn (1953) asserts that teacher behavior changes on a daily basis and observers have to rate teachers over a period of time and on many separate occassions. Flanders (1964, 1968) states that no matter how reliable or valid the observer's assessment procedure, the results are somewhat distorted, for the teacher tends to put on an act while being observed. Operating, similarly, is what is called "demand characteristics." Accroding to Gephart and Antonoplos (1969), Orne (1962), and Rosenthal and Jacobson (1968), the subject (in this case the teacher) is willing to cooperate as indicated by the fact he has usually consented to the observer's presence, and therefore perceives an "acceptable" role, which, in turn, changes his behavior. The presence of the investigator (in this case the observer) creates what is called the "hawthorne effect" - novelty, awareness of participation, and/or an altered situation - on the subjects and who the subjects are interacting with (teachers and students), according to Bloom (1969), Cook (1962), Gephart and Antonoplos (1969), and Rosenthal and Jacobson (1968), and these effects are to complex to determine (Gephart & Antonoplos; 1969; Masling & Stern, 1969). According to Gephart and Antonoplos (1969), Rambo (1969), and Rosenthal and Jacobson (1968), the investigator (again, the observer) transmits what is termed "bias effect" to the subjects and who they are interacting with (again, the teachers and students), that is, his own hunches or prejudices which are often one of the factors which prompted the study, in a way to alter the subject(s)' behavior.

Sorenson, Husek, and Yu (1963), Ryans (1952), B.C. Smith (1967), and Sprinthall, Whiteley, and Mosher (1966) contend that observers are influenced by their own values and role interpretation of what is a "good" teacher. Jenkins (1960) and Brown (1967) feel that even the age and sex of the observer and teacher influence the rating. Brown (1967) and Medley and Mitzel (1958) are

of the opinion that cues upon which the observer bases his judgments vary in importance from observer to observer, even with the same observer, for different teachers. Guilford (1954) and Ryans (1952) maintain that the observer is subject to the "halo effect," whereby the observer rates the teacher's behavior in the direction of the general impression of the teacher. Guilford affirms that the observer's ratings are also distorted by the following: (1) "error of leniency," a tendency of the rater (in this case the observer) to rate low or high, no matter what the reason; (2) "error of central tendency," whereby the rater (observer) is reluctant to make extreme judgments about others (teachers); (3) "constant error," whereby the rater tends to rate others in the opposite direction of his own behavior - for example, the observer who is businesslike tends to perceive the teacher as less businesslike, or the observer who is not too businesslike tends to perceive the teacher as more businesslike. (On the hand, Guilford refers to Hollingworth (1922) who claims that for "good" teacher behaviors there is a positive relationship between the observer's possession and ability to evaluate it, and vice-- versa.) Flanders (1960) believes that each of us has a preferred set of teacher behaviors, and even though the observer has a specific list to interpret, he tends to concentrate on the favored items and by passes the others.

Biddle (1964) alleges that the observers not only are biased, but they lack real knowledge concerning the specific classroom problems, which are affecting the teacher's behavior. Jenkins (1960) believes that the same teacher behavior means different things to the students than to the observers. Symonds (1955) claims that observing teacher behavior is of little value, for the basic referents of effective teaching are linked with the teacher's personality.

---

* Actually, Guilford is specifically referring to raters, not obersvers; however, in effect, the observers are rating teacher behavior with some kind of rating scale. Thus Guilford's discussion is germane to observers.

Biddle (1967) and Guba and Getzels (1955) point out the problem of "observer loading," namely, it is "humanly impossible" to objectively observe or just observe, all teacher behavior or classroom phenomena, and it is unlikely, too, according to Guba and Getzels (1955) and Guilford (1954) that the phenomena which are observed will be given appropriate weights. Guilford maintains that the observer cannot rate all teachers equally well on all traits. Rosencranz and Biddle (1964) and Walberg (1969) aver that only overt behavior is measured by the observer. Galloway (1966), Gibb (1960), and Ornstein (1969a) assert that nonverbal and/or intangible entities pertinent to teaching are often overlooked by the observer, i. e., an expression or glance that is easily understood by the students. Torrance (1960) states that even though the teachers says the "right words" or behaves in the "right way," his "real" attitude is evident to the students [but often overlooked by the observer], which, in turn, affects the classroom process. Gibb (1960) claims that the direction of verbal communication - who talks to whom, if the teacher's statement is directed at an individual student or at an individual student as a member of a clique or class, is important but difficult for the observer to discern.

Sorenson et al (1963) contend that most observers cannot state precisely the reasons for their judgments. B. O. Smith (1967) believes observers are unable to observe teachers systematically. Guilford (1954) points out that raters (observers) often lack sufficient time to make their evaluations. Turner (1964) showed that the observer's intelligence accounts for as much as 15 percent of the score variance of a measured teacher behavior. Walberg (1969) contends that as much as 20 percent of the score variance of a

specific teacher behavior cannot be objectively assessed. Flanders (1965) and Ryans (1960), in connection with their own studies, admited that their trained observers had difficulty in distinguishing one teacher behavior from another as belonging to a specific class of behaviors. Guilford (1954) and Ryans (1964) assert that once observers have been trained, it cannot be assured that their reliability will remain high over a period of time.

Although it might be pointed out that some of the above problems can be reduced with the introduction of visual and/or auditory tapes, according to Biddle (1967), the noise level of the class, the mechanical problems, and the cost of recording do not make the mechanized approach as valuable at it might seem. Also, Biddle points out that the recordings are produced by the observer or filtered through his eyes, and therefore, accroding to this author, they still incorporate and reflect most of the above contaminating factors.

Student behavior and achievement. Barr (1950) and Cogan (1968) are of the opinion that practical methods for evaluating teacher as a function of student behavior have not yet been developed. A major problem seems to be that different student behaviors are assessed with different teacher behaviors, making it difficult to obtain a consistent thread or relationship. Three examples should indicate the infinite number of combinations. Ryans (1960) assessed 4 student behaviors (alertness, responsibility, confidence, and initiative) with 18 teacher behaviors or 3 broad teacher patterns (warm, understanding, friendly; responsible, businesslike, systematic; and stimulating, imaginative, surgent). Perkins (1964) assessed 9 student behaviors which, for the sake of brevity, may be subsummed under two categories (work activity and social activity) and 10 teacher behaviors - along a supportive-

nonsupportive continuum. Harvey, Prather, White, and Hoffmeister (1968) assessed 7 student behaviors (cooperativeness, involvement, activity, nurturance, seeking, achievement*, helplessness, and concreteness of resonses) with 3 teacher behaviors (resourcefulness, dictatorialness, and punitiveness.) It should be pointed that not one student or teacher behavior is the same, and assuming if there was one, the definitions would probably differ. In short, dissimilarity of student and teacher behaviors, along with differences in definitions, make comparisons extremely difficult and often misleading. To the author's knowledge, no statistical formula could accurately analyze together the different behaviors, only show loading or correlational effects.

Student behavior seems functional to countless other variables, which are often uncontrollable and too multidimensional to analyze effectively.[8] A few examples suffice: subject matter (Gump, 1964; Ryans, 1964), peer-group relations (Eisner, 1963; Gump, 1964; Jensen, 1955), classroom activities Biddle, 1964; Gump, 1964), school conditions or school norms (J. M. Jackson, 1960; Getzels & Thelen, 1960; Lehmann, 1960), and community relations (Biddle, 1964; Jensen, 1955), to name a few.

Examining the subtleties of just one variable - subject matter - with regard to student behavior should make the reader more aware of the immense problem involved with coping with all the known variables for puposes of evaluating teacher behavior. Eisner (1963) points out that some subjects call for energetic, active student behavior, i. e., music, drama, and physical education; others usually demand a monotonous, quiet, drill-like atmosphere, i. e., mathematics and foreign langauge. The question is, does the investigator realize, consider, and adjust his analysis to the different

---

* The reader should note that "achievement" is not an example of student behavior.

[8] Infra, pp. 28-29.

reasons for the different atmospheres. To the knowledge of this author, the answer seems to be no; therefore, the research on teacher behavior tends to show that mathematics and foreign language teachers are more authoritarian, businesslike, and/or responsible than music and English teachers, and no qualifying explanation is provided.

Assuming the investigator is aware of the many variables, how does he evaluate them regard to student behavior. For example, the teacher asks a question, but no answers are forthcoming. Is this because the observer is in the room? Is it because the students are bored or confused? How many students are hungry, or haven't had breakfast? Are the students reacting to their present teacher or their pevious one? Is the weather or lighting* influencing the students? Does the day of the week or time of the day influence the students? Is student behavior the same on Monday morning and Friday afternoon? Is there a basketball game scheduled for the afternoon. No research seems to consider these subtle factors in their assessment of student or teacher behavior, and to dismiss them as being insignifact or minor is mistaken, for they total a significant part of the classroom situation.

---

* Recently, an experienced teacher argued that if two elementary-school classes, A and B, were matched according to I.Q. and reading achievement, then pre- and post-tested after one year, if the students in class A were reading one year higher than the students in class B, it would be safe to assume that teacher A was a "better" teacher. The author remarked that there were still too many variables, many unidentifiable or too nebulous, to warrant such a conclusion. The students in class A could have been in a room for the whole year in which the sum shined through the windows, whereas the students in class B might have been situated on the dark side of the school. Would the investigator, teacher, or examiner note the difference? Do we look for such small, subtle differences? How important are these differences? No one seems to mention such variables, which does not necessarily mean they are irrelevant. Perhaps the amount of light or the difference in the amount of light in a classroom is a key factor, which we tend to ignore. We don't really know; there are no studies, it seems, that say otherwise. The teacher contended the author was being absurd. "No, I'm reflecting the absurdity of research on teacher behavior." Total all the so-called minor and unidentifiable variables, include, too, the variables we recognize, but cannot agree upon with regard to defintion and weighting, and we have no clear conclusions.

Gump (1967), J. M. Jackson (1960), and Medley and Mitzel (1963) contend that observers have difficulty in distinguishing between teacher-behavior intent and effect on student behavior. Atkin (1967-68) contends that there is no agreement on what constitutes desirable student behavior. By the same token, it is possible for the teacher to know the rules of"good"behavior or to implement "good" teacher behavior, assuming we could agree on what "good" teacher behavior is, but this does not necessarily guarantee "good" student behavior. According to Eisner (1963), James (1958), and Ornstein (1969a), there is an intangible relationship between teacher and students - which affects students' behavior - but cannot be prescribed or defined.

Flanders (1960), Leacock (1969) and Keislar and McNeil (1957) contend that teacher behavior is determinant of student behavior, that is, teachers adopt a particular behavior to the extent they perceive student interest and rapport. On the other hand, Yee (1968) argues that student behavior is more responsive to teacher behavior. Gump (1964) and Packer and Packer (1959) take a middle position, that both teachers and students interact to the extent that it is difficult to determine who is reacting to whom, or which is the independent variable. If this is the case, that investigators cannot agree upon whose behavior is independent and dependent, correlational and regression analysis may not be a "wise" formula to use, since it is difficult to designate the predictors from that which is predicted, the symbol X from the symbol Y, which is important to know.

One might question how the observer observes the students' behavior. If the observer is in the rear of the room, which tends to be the usual method, he perceives the back and external part of the cranium, not the students' face and telling gestures. Surely, the observer must miss some behavior, perhaps, even be decieved. If the observer positions himself on the side or front of the room, the "hawthorne effect" and "observer's biases" are probably enhanced.[9]

---

[9] Supra, p. 13.

Finally, since the assessment of student behavior usually involves the observation of students, the problems of observations are generally applicable for assessing student behavior.[10]

Limitations are apparent in using student achievement as a criterion for assessing teacher behavior. Cronbach (1966) and P. W, Jackson (1968) contend that learning principles are vague in relation to the actual classroom process. Scandura (1966) point outs that student achievement is often incidental to "good" teacher behavior. Glazer (1965), Skinner (1965), and Stolurow (1965) assert that student achievement is inconsistent with "good" teacher behavior.

Cronbach (1963), French (1961), and Justiz (1969) allege that it is difficult to equate the effects of a particular teacher with student achivement. B. O. Smith (1967) declares that is is difficult to distinguish which modes of teacher behavior - the verbal interchange, general strategies, reinforcement techniques, etc. - are related to student behavior. Bloom (1964), J. M. Hunt (1961, 1964), and Kohlberg (1968) point out that environmental factors influence student achievement. Hedges and MacDougall (1964) and Justiz (1969) claim that the variations in student personality, intelligence, past achievement, and/or the above environmental factors make it difficult to measure objectively student achievement as a function of teacher behavior. Anderson (1954) mentions other contaminating factors, such as mass media, low or high pretest scores, time interval between the test- and retest, school conditions, etc. Biddle (1967) points out that the initial and final achivement tests are usually administered in a relatively short time interval; therefore, according to this author, the magnitude of differences between tests tend to be small. In this connection, Bloom (1964) and McNemar (1958)

---

[10] Supra, pp. 12-16.

claim that it is difficult to obtain reliable scores when the magnitude of differences is small.

Ackerman (1954), Anderson (1954), and Orleans, Clarke, Ostreicher, and Standlee (1952), and Symonds (1955) maintain that achievement tests only measure a small portion of the desired change expected of students as a result of the teaching-learning process. It omits, for example, the personal and social growth of the student, the developmental tasks outlined by Erikson (1950), and Havighurst (1950).

For the greater part, most educators will admit that achievement tests often lack acceptable reliability and validity scores[11] In this connection, achievement tests are considered to be culturally biased (Anastasi, 1967; Davis, 1948; Deutsch, Fishman, Kogan, North, & Whiteman, 1964; Manning, 1968), and tend to discriminate against creative and/or intelligent students (Anastasi, 1967; Hoffman, 1964); they often lead to erroneous connotations (Anastasi, 1967; B. M. Smith, 1967) and negative "dysfunctional" outcomes such as the "self-fullfilling prophecy" (Clark, 1965; Rosenthal & Jacobson, 1968).

Finally, if an investigator or an observer is in the classroom, observing teacher behavior, or especially if he is administering a student achievement test, many of the problems concerning the "hawthorne effect" and "bias effect" appear to operate with the students, too.[12]

Ratings based on recall. Teacher-behavior ratings based on recall are generally made by supervisors, teachers, and/or students.* Studies of what constitutes "good" teacher behavior, by Anderson (1954), Borg (1957), Crawford

---

[11] The reader is advised to consult Buros (1965), Cronbach (1959), Guilford (1954), and Hoffman (1964).

[12] Supra, p. 13.

* Observers fall into this category, too, but they are not discussed here, since they have been previously mentioned.

and Bradshaw (1968), and Getzels and Jackson (1963) show that while evaluations made by supervisors, teachers, and students may be consistent they are often contradictory, or not significantly related. Similarly, Barr (1946) and Veldman and Peck found marked disagreement between supervisors' and students' evaluation of teacher behavior. Willard (1957) showed there was a lack of relationship between supervisors' behavior reports of teachers and teachers' self-reports. Symonds (1955) found that peer ratings of teacher behavior and students' rating of teacher behavior lacked a relationship.

With regard to supervisors, Stern (1963) contends that supervisors' ratings of teacher behavior are influenced by "factors which are irrelevant" to effective teacher behavior [p. 421]." Start (1968) showed that teacher personality profiles similar to their supervisors' personality had the highest rating for teacher ability. Hawkins and Stoop (1966) showed marked contradictions with supervisors' ratings of teacher behavior and subsequent evaluation reports made by the same supervisors.

With regard to teachers, Tschechtelin (1953) found that teachers tend to overrate their colleagues. McCallon (1966) found a positive correlation between teachers' rating of self-behavior and their rating of students as desirable to teach. T. E. Smith (1965) showed a wide discrepancy between the self-evaluation of young and old teachers.

With regard to students, there tends to be substantial agreement that they are the most worthwhile and honest raters of teacher behavior; moreover, they appear to be reliable raters - with coefficient scores reaching the .90's (Paraskevopoulous , 1968; Symonds, 1955; and Veldman & Peck, 1963), even after a one-year interval (Christensen, 1960). Remmers (1963) affirms that as long as 25 or more students' rating of teacher behavior are used there is

considerable reliability. Beyond their assumed reliability, the students are considered more valuable as raters because they see the teacher perform on many occassions under varied conditions, according to Paraskevopoulous (1968) and Veldman and Peck (1963). Christensen (1960) mentions that students within the same class can be regarded as many observers rating one teacher. H. C. Hunt (1942) feels they are the best judges because as a group they represent a constant variable. Cogan (1968), McNeil (1967), and Paraskevopoulous (1968) are of the opinion that since the students' feelings are a major factor in determing the classroom climate, they are most qualified to rate teacher behavior. Gage, Runkel, and Chatterjee (1960) showed that teachers, given students' ratings of their "ideal" teacher, changed their behavior in the direction of their students' "ideal" teacher, as indicated by students' subsequent assessments of their teachers.[13]

Nevertheless, investigators have pointed out that teacher-behavior ratings which are based on student recall do have limitations. Bryan (1941)[9] contends that students lack knowledge of what is "good" teacher behavior; they are immature, their ratings are influenced by how easy the teacher is, and their ratings negatively affect teacher morale. Rees (1969) alleges that teachers affect students in different ways, and what accounts for these differences is not so much the teacher's behavior but the students' personalities. In this connection, then, the teacher can employ "good" teacher behavior but be rated as a "poor" teacher, because the students' rating reflect their attitudes and values. Remmers (1963) and White and Anderson (1967) showed that perception of teacher behavior varies with student-achievement level. Drawthorne (1954) showed that as interaction between students and teachers increase, student ratings of teacher behavior tend to improve. Beck (1967)

---

13 Supra, p. 19.

points out that student ratings, while reliable, vary from grade to grade. Walberg (1969) contends that low student-reliability scores for specific teacher-behavior items may be ignored so long as the broad teacher-behavior dimension averages out to be acceptable. Yet Cronbach (1959, 1963) asserts that item data are more important than the total score because the latter conceals judgments about the importance the importance of specific data and is less suitable for further improvement of the instrument.

Referring, now, to the human rater in general, no matter if he is a supervisor, teacher, student, etc., the problems listed by Guilford (1954) under the discussion of observation - "halo effect," "error of leniency," "constant error," "error of central tendency," - tend to affect raters who are assessing teacher behavior on recall, too.[14] Other factors that tend to affect raters, according to Guilford are: (1) sex, (2) age, (3) intelligence, (4) understanding of directions, (5) understanding of purposes, (6) sufficient time to complete the ratings, (7) possession of the traits being measured, (8) different criteria raters employ for assessing the same trait or teacher behavior.[15]

According to Crowne and Marlow (1960) and Edwards and Diers (1962), when dealing with items about personality or behavior, raters often give answers they perceive as right to the investigator or text examiner. Rambo (1969) contends that if the test examiner is not perceived as a member of the raters' "reference group," they tend to give "socially acceptable answers." Maccoby and Maccoby (1954) and Rambo (1969) mention such contaminating factors as the way the raters perceive the test examiner's clothing, socio-economic class, race, and name. Allport (1935) maintains that attitudes are difficult to measure, because raters have two different attitudes - one for friends and relatives, the other for formal surveys. Biddle (1964) claims that raters are not always motivated or honest, and that findings often reflect their lack of information concerning the "desirability" or "undesirability" of

[14] Supra, p. 14.

[15] The 8 problems listed above are representative of 33 common problems which affect raters and ratings, summarized at the end of Guilford's chapter entitled, "Rating Scales," pp. 263-301.

what is being measured. Remmers (1963) maintains that raters, because they are human, are "imperfectly reliable" and their judgments are not "highly valid [p. 372]," being suspectable to "selective perception, memory, and forgetting," as well as "lack of sensitivity to what may be ... important [p. 329]." Berde (1969) points out that test interpretation of teacher behavior vary according to raters; moreover, according to Himmelfarb (1969), the scale values (assuming the investigator weights each item) are determined by judges or raters judging the location of each response in terms of "desirability," "favorability," "importance," etc., and their attitudes are biased, thus influencing (according to this author) the subsequent scoring of each respondent.

Finally, the problem of what is acceptable reliability and validity seems germane to the construction of rating scales.[16] In this vein, Biddle (1964) calls for the elimination of rating scales until the problems are better understood and controlled.

Personality tests. According to Getzels and Jackson (1963), psychologists are unable to agree upon the definition of personality or specific personality traits; therefore, it seems that data provided by one instrument do not necessarily yield analogous data, even though the findings may correspond. Phillips (1967) and Symonds and Dudek (1956) claim that psychological

[16] See Buros (1965), Cronbach (1959), Guilford (1954), and Remmers (1963). Also, see pp. 4, 12, 15, 18, 24, 26.

tests are restricted because we do not know which personality traits are predictive of "good" teacher behavior. Michaelis (1954) contends that not enough is known about personality traits to predict teacher behavior. On the other hand, G. B. Johnson (1957) affirms that psychological tests are limited because teacher behaviors are inadequately defined. Getzels and Jackson (1963) conclude that "very little is known for certain about the nature of teacher personality, or about the relation between personality and teacher effectiveness [p. 574]."

Kerlinger and Kaya (1959) and Taylor (1968) believe that most psychological tests have uncertain validity. Budd and Blakely (1958) show that psychological tests administered to teachers can be biased in favor of extreme responses. Sorenson (1956) points out that signing the answer sheet of a psychological test will make a significant (.05 level) difference in the teachers' answers. Callis (1950), Coleman (1954) and Medley (1961) show that psychological tests are susceptible to "faking" by teachers. As Medley (1961) indicates, "You can't believe the answers teachers give; [moreover], those who know how to get along with pupils also know how to get along on personality tests as long as they are not too subtle [p. 153]."

Taylor (1968) contends that it is inappropriate to rate someone on some point along some continuum because personality traits are abstract. Gordon (1965), Guilford (1954), and Cronbach (1959) are of the opinion that "unsure" or "neutral" responses on personality or attitude tests tend to make the instrument less valid. On the other hand, Getzels and Jackson (1963) and Turner (1968) contend that "absolute" responses often force the respondent to make an unwarranted decision. Walberg (1967) points out that it is impossible to control all the teacher-behavior variables - sex, intelligence,

age, education, experience, etc. - and correlate psychological scores and teacher behavior.

Strong (1943) affirms that personality tests are usually given to a group of teachers, without distinguishing differences in subject field, grade level, education, etc., thus averaging out differences that may exist. Wehling and Charters (1969) claim that most psychological tests are evaluated in terms of teacher success or teacher behavior, as if an "ideal" existed. Getzels and Jackson (193) maintain that teacher personality scores are over-simplified, and do not really help the descibe the effect of the teacher's personality upon the student or class, the teacher-student interaction - what teaching involves.

Finally, according to Getzels and Jackson (1963) teacher-personality scores tend to be useless, contradictory, and lacking in psychological and "common sense." Also, many of the other problems of reliability and validity, discussed earlier, seem germane to personality tests.[16a]

Some Additional Problems Related to Research on Teacher Behavior

Rosencranz and Biddle (1964) point out that much of the previous research on teacher behavior is nonculmulative, in the sense that, according to AERC (1952), Biddle (1964, 1967), Gage and Unruch (1967), and Medley and Mitzel (1963), researchers measure different phenomena and variables, use different terms, methods, and assumptions. Bloom (1964) contends that when findings are not in harmony with existing data, it behooves the researcher to explore further into his results and examine the reason; however, Biddle (1967) and Gage and Unruch (1967) are of the opinion that ideas and findings on teacher-behavior research are usually promulgated without much reference - and with apparent disregard - of what others say or report. Broudy (1969), Gage (1968),

---

16a See p. 25, fn. # 16.

and Rosencranz and Biddle (1964) assert that researchers are lacking a common framework to work with.

The problem of distinguishing, controlling, and analyzing what seems to be an endless amount of variables may be too difficult, if not impossible, for obtaining worthwhile data.[17] Accrording to AERC (1952), Biddle and Ellena (1964), Gage (1968), Hearn (1953), Ryans (1960, 1964), and B. O. Smith (1967), there is simply no adequate criterion and/or list of variables against which a list of teacher behaviors can be validated and/or compared. Broudy (1969) and Flanders (1964) allege that not all the variables relevant to a situation are known.[18] P. W. Jackson (1968) affirms that classroom events occur at such a rapid pace, involving 200-300 interpersonal changes per hour; it cannot be accurately systematized into a scheme that can help teachers in their actual situation. On the other hand, Atkin (1967-68), as

---

17 A few examples of teacher behavior variables are the time, place, school morale, school goals, teacher training, sex, age, grade level, type of classroom, community, etc. Combine this list with an endless list of student variables (see p. 17, for a few), as well as unidentifiable variables (see p. 18).

At best, the findings of a study on teacher behavior should be considered relative the variables being manipulated - bearing in mind that many haven't been identified. For example, does a "friendly" teacher, which Getzels & Jackson (1963) contend "obviously" connotes a "good" teacher (supra, p. 9), have the same effect all the time, in all schools, no matter what his age or sex, no matter what grade level or subject, in the classroom as well as when coducting student traffic in the cafeteria, hallways or auditorium, and with all types of students? Are there shades of differences or major differences, and to what extent, with which variables?

For a discussion of several variables, the reader is referred to AERC (1952), Biddle (1964), Flanders (1960, 1965), Gump (1964), Getzels & Thelen (1960), Peterson (1964), and Ryans (1960, 1964).

18 Supra, p. 18.

previously mentioned, contends that insignificant variables are often manipulated into research, producing trivia.[18a] French (1961) affirms that many variables are unpredictable. Parker and Parker (1959) believe that individual variables are "coupled" - meaning that each affect the other, combining into new components; moreover, this author contends that the new components affect other variables and cause, still, other components, some of which cannot be thoroughly distinguished or measured. Similarly, French (1961) contends that tow or more variables do not necessarily reflect casual relationships, or reflect what they seem to show, but may reflect other variables which act upon the ones that seem more evident. Also, Ebel (1967) and French (1961) assert that variables are mulidimensional, not linear, and therefore, confusing and difficult to assess. In short, there are an infinite number of variables - whose interaction and importance are relative; moreover, some of which are unknown, uncommon, unique, and unpredictable, but important although we are not sure to what extent - yielding uncontrollable data, making assessment of teacher behavior difficult, if not impossible and worthless.(Ornstein, 1970a).

Analysis of teacher-behavior research may be beyond scientific analysis, because the act of teaching, itself, may be unscientific. Ebel (1967) claims that the process of teaching is not a natural phenomeon that is suitable or controllable for scientific inquiry. P. W. Jackson (1968) states that teacher behavior is difficult to assess because the act of teacing involves working with complex organisms. Flanders (1964) and French (1961) maintain that teaching is novel, not absolute in the sense that all teacher behavior and teaching situations are new, making it obscure for researchers. As previously

---

18a
Supra, p. 4.

mentioned. P. W. Jackson (1968) maintains that teaching is too complex for an abstract or scientific description.[19] Eisner (1963), Klein (1969), and Macdonald and Zaret (1968) aver that teaching cannot be quantified into global or recognizable terms. Travers (1966) believes that a technical language has not yet been developed, one that is empirically based, to evaluate the actual teaching phenemona. Gibb (1960), Galloway (1966), Macdonald and Zaret (1968), Ornstein (1969a), and Walberg (1969), as previously mentioned, assert that many teaching acts, especially non-verbal ones, go unnoticed, or are difficult to make sense out of and evaluate.[20] Klein (1969) contends that empirical data cannot be used to measure teacher behavior because teaching is transcendental - obscure, incomprehensible.

Atkin (1967-68) affirms that researchers are using, for their analysis of teacher behavior and teaching, behavioral-science approaches, not educational theory or approaches applicable for classroom analysis. James (1958) affirms that psychology is a science and teaching is an art, and psychological approaches [or behavioral approaches, as suggested by Atkin] cannot fully describe an art. Eisner (1963) and Flanders (1964) contend that teaching is both a science and an art, and cannot be completely analyzed by scientific methods. James (1958) claims that a science can indicate rules which an act in this case teacher behavior can follow, but the specfic act is an individual adaptation within the defined boundaries, producing an ardous research situation. Bloom (1969), Eisner (1963), P. W. Jackson (1968), and Wall (1969) maintain that teaching depends on feelings, hunches, and/or insights, and they are often more useful "to determine what and how" to teach than scientific findings.

---

[19] Supra, p. 4.

[20] Supra, p. 15.

The fact that teachers generally seek practical, "cook-book" approaches (Jackson, 1968; Ornstein, 1969c; Sceviakov & Redl, 1956; Wilkerson, 1966), although often condemned by educators as mechanical or hazardous, suggests that teacher-behavior research is meaningless for teachers; it may suggest that scientific analysis and/or theoretical formulations are inapplicable to the classroom, too. Eisner (1963), French (1961), Travers (1964), and Turner (1964) claim that while teacher-behavior research can be formulated, it is unsuitable to practice, to the actual classroom situation because every teaching situation is different. Jenkins (1960) maintains that the teacher who attempts to apply research findings to his classroom may not obtain similar or expected results, even though the teacher carries out the same behaviors. Ornstein (1969c) and Scheviakov and Redle (1956) contend that the best advice [which may be based on research] can sometimes be harmful, because each teacher, student, and group of students - each situation - varies.

French (1961) affirms that the researcher often fails to put himself in the position of the teacher, and therefore, omits relevant facets of teacher behavior. According to P. W. Jackson (1968) and Klein (1969), teachers use vague terminology to define their own classroom behavior and are unable to systematize or explain what they did. Likewise, Rivlin (1965) and Wilkerson (1966) point out that many teachers know what they are doing in the classroom, but are unable to specifically state it into precise terms. P. W. Jackson (1968) maintains that teachers and researchers use different terms to describe the same teacher behavior; moreover, the researchers among themselves use different terms.[21]

---

[21] Supra, pp. 9-11.

D. W. Johnson (1969) maintains that merely to inform teachers about what constitutes "good" teacher behavior does not necessarily mean that teachers will change or even assume that is desirable. Strong (1943) contends that teachers are conservative and are likely to resist changing their own behavior.[22] Similarly, P. W. Jackson (1968) believes that teachers lack the time to concentrate on modifying their teacher behavior.

Even when teachers seek feedback from researchers, especially from doctorate students who are conducting a study to complete their thesis, often, they are not provided with such research, even though the findings can be stenciled and mailed to the principal or individual teachers. Many teachers lack understanding of research techniques (Ornstein, 1969b), and they are unable to interrupt findings, even if the researcher provides them with data. Other teachers seem to dismiss or resist research with "that all good theory, but it doesn't work [Ornstein, 1969c, p. 84]."

Many teachers seem no longer willing to cooperate with investigators from the colleges and universities, because some of them or their partisans, since Sputnick, have an increasing, near-compulsive disposition to criticize teachers. Since the War on Poverty, this criticism seems to be focused on teachers of the disadvantaged (Ornstein, 1970b). The criticism seems unfair, wholesale, and flagrant (Ornstein, 1967, 1968), often couched in the angry rhetoric of angry exaggeration (Ornstein, 1971), generally a biased presentation about one ghetto school - sometimes a few - whereby the uncritical reader tends to make generalizations about all ghetto schools and ghetto teachers. (Ornstein, 1970b).

Many of these critics are divorced from the colleges, but in response to

---

[22] Supra, p. 23. See Gage et al, what seems to be an opposite viewpoint.

their widespread criticism, often based on emotionalism, rhetoric, or phoney "liberalism"(Ornstein, 1971), school teachers and administrators tend to generalize, too, that most outsiders are potential critics and a possible threat to the school or school system. Writers like Friendenbrg, Hentoff, Herndon, Holt, Kohl, Kozol, and Schragg, who gush over a review each other's books as if there they are a part of an organized conspiracy (Ornstein, 1971), tend to be popular with preservice and younger teachers, [as well as black power advocates], because their views often correspond with the anti-establishment trend which is sweeping across the country (Carras, 1969). According to Havighurst (1968), and Ornstein (1971), these critics tend to be dangerous and irresponsible, for they call for the destruction of the schools without offering a viable alternate.

As a reaction to the above writers, teachers and administrators seem no longer welcoming researchers from the colleges and universities. The lines of communication between school personnel and researchers seem strained, and this has serious implications for doctorate students who are often dependent on the teachers' and school officials' goodwill. Whether teachers are anti-research is no longer the number-one problem; they seem [illegible] anti-researcher. The problem seems to be compounded by growing tension in the inner-city schools, whereby a particular school has too many problems to permit an outsider or researcher to perceive, perhaps even write about, the tension. Similarly, in context with the black power movement, northern schools that are controlled, in part, by black communities or which have a large and/or militant black staff no longer welcome white investigators.* This has serious implica-

---

* A great deal of the unverified statements which appear above are based on conversations with several teachers, principals, and superintendents, from Chicago, Philadelphia, and New York City, as well as observations of many inner-city schools. For purposes of professional courtesy, the people and schools remain anonymous.

for all types of educational research which are related to the schools.

Conclusion

Lameke's (1955) comment of more than a decade ago bears repeating:

> If the research during the last three years were to be wiped out in the field of medicine, agriculture, physics, or chemistry, our lives would be materially changed. If research in the area of teacher personnel during the last three years would vanish, education and educators would continue as usual [p. 192].

It is sad but true that the most serious researchers on teacher behavior, a closely related field to teacher personnel, would probably not only make the same statement, but they might go back to the turn of the century.[23]

Ebel's (1967) criticism of research in general seems relevant to research on teacher behavior:

> Even today, when the prestige of science is at its height, most ... of the knowledge with which we guide our lives and solve our problems has come, not from controlled experiments, but from practical experience p. .
>
> Most of the problems we face in the world today, including our educational problems, involve questions of purposes and values ... decisions that science could not possibly make for us p. .

Most of the knowledge and acts which guide the teachers' behavior in the classroom is based, not on research, but personality, "common sense," and experience.[24] Teaching involves an on-going interaction between teacher and students, problems arise that must be dealt with on the spot, as it occurs - research does little good at that moment,[25] every situation is somewhat different,[26] feelings, insights, "common sense," etc. seem more important.

---

23 Supra, pp. 1-2.

24 Supra, p. 7.

25 Supra, p. 31.

26 Supra, p. 30.

Concerning the limitations of theory in practical realms, Eisner (1963) cites Aristotle:

> ... it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits ...[

If this viewpoint is translated to research on teacher behavior, this may suggest why the field is beset with so many problems and why little worthwhile data have been forthcoming.[27] The practical aspect of teacher behavior and teaching may just be undefinable by research, or unable to be subsummed by a research principle. Assuming this may be true, and the problems seem to indicate this,[28] how long will it take researchers to admit it?

Finally, many readers may find themselves agreeing with many of the above limitations and recommendations below of research on teacher behavior; they should note, however, the content of the paper is subjective and "arm-chair" in nature. No critical evaluation of the many studies and references has been attempted, due to space limitations, and the interested reader - one who wishes to make his own approximate resolutions - is forced to read the materials on his own.

Recommendation for Future Research on Teacher Behavior

1. Granted, the status of research on teacher behavior is flaccid and imperfect; however, it can be improved - to a point, however, which is unclear.
2. Rather than being trapped in analyzing teacher behavior, the invesitigators should first spend time in understanding its depth and complexity.
3. Investigators need to agree on (a) operational terms, (b) content of

---

[27] Supra, pp. 1-2.

[28] Supra, pp. 1-32.

inventories, and (c) measurements of teacher behavior.

4a. Teacher-behavior terms should be formally defined, as are words in a dictionary (Carroll, 1964).

4b. Teacher-behavior items should be neutral, that is, should take the same form regardless of subject, grade level, etc. (Meux & Smith, 1964).

5. There is need to refine and formulate agreed-upon teacher behavior inventories. Terms and categories should be established in view of validity - content, concurrent, and predictive.

6. There should be agreement on measurement instruments, and on which instruments have equality of weights and units, beginning at the same point and preferably at zero.

7. Assumptions for using parametric tests, which are most popular among the researchers, should be indicated or at least made clearer. In some cases, actually, nonparametric tests should have been employed; they were not, therefore, the findings are distorted.

8a. More attention is needed to understand the nature of teaching and the classroom process. Researchers tend to interpret data in terms of behavioral sciences. There is the need to translate findings into terms that are both (a) applicable to teaching and the classroom process, as well as (b) comprehensible for teachers.

8b Under the guise of good scholarship, professors and researchers, especially in the field of education, tend to write for the benefit of their colleagues, or at their audience level, partially as if the readers know about what is being promulgated, but had failed to comprehend it. This may be beneficial, but it does not help or affect most of the people who could benefit from "new" knowledge, in this case teachers.

8c. In the meantime, teachers should try to understand research and translate findings to their own classroom situations.

9. Teaching involves a teacher-student interaction. Teacher-behavior research should be formulated in relationship to both teacher and student behavior. One without the other is relatively misleading and useless.

10. There is a need to learn to what extent teacher behavior is a function of personality (Getzels & Jackson, 1963).

11. In analyzing teacher behavior, it is important to consider the context in which it occurs. Much of the research, now, tends to treat teacher behavior as an isolated entity.

12. There is a need to control variables, at least the major ones, and try to make sense under what condition what teacher behavior is desirable to what extent.

13a. There is a need to bring together and synthesize the mumerous criteria, as well as the concepts and methods of conducting teacher-behavior research, into a framework which consists of a critical examination and comparisons, so that additional data can be hypothesized, developed, and analyzed in terms of previous data.

13b. Preference should be given to teacher-behavior criteria that have been commonly studied to maximize chances of validity and subsequent comparisons of data.

13c. Variables that cannot be agreed upon or organized should be tentatively discarded, for it would be easier, then, to validate teacher-behavior criteria, compare findings, and formulate hypotheses and theories.

14. Teacher behavior and teaching are often described as abstract, nebulous processes. There is a need to break these processes down into smaller and concrete components, which are recognizable and agreed upon, for purposes of

analysis.

15a. A micro-analytical approach for the study of teacher behavior may be helpful, with well-defined criteria and agreed upon definitions, so that data may be more controllable, objective, and useful.

15b. Several micro-analytical studies might lead to teacher-behavior theories (Gage, 1968).

16. To the author's knowledge, very little, if any, research on teacher behavior has been conducted when teacher and students were using technological hardware in the classroom or school. Educational technology is increasinly used by teachers. Teacher-behavior research should focus in this direction, since this seems to be a future trend with classroom learning.

17. Investigators should take more advantaged of educational technology (computers, videtapes, records, etc.,) for purposes of facilitating, improving, and analyzing their research.

18. Much teacher-behavior research is conducted by doctorate students for purposes of earning a higher degree. University regulations, coupled with the candidate's desire to complete the study within approximately one year and his committee's desire to see him complete it, tend to make the candidate "play it safe" with a trivia problem. The idea is to get one's degree and then contribute something worthwhile to the field. Similarly, doctorate candidates usually lack one or more of the following: sufficient time, financial aid,, staff assistance or manpower, expertise, equipment, facilities, etc. This often leads to a somewhat worthless or useless study. For this reason, doctorate candidates should no longer be encouraged to conduct research on teacher behavior, unless it is a part of a more comprehensive study with sufficient funds, directed by an authority in the field.

19. Investigators seem content to conduct their studies wherever they can find subjects (Getzels & Jackson, 1963), especially doctorate candidates. Teachers and students should be chosen by systematic selection and sampling, assuming they will cooperate, in order to make comparative studies.

20. Longitudinal teacher-behavior studies should be conducted with agreement on criteria and variables.

21. Periodically, a nationwide teacher-behavior study might be conducted, noting socio-racial-geographical differences along with other agreed-upon variables, as a means for comparing other studies and teacher training programs.

22. Research on teacher behavior should be conducted in relation to the effects of various teacher-training programs.

23. We do not know how to train "good" teachers; we rely on descriptions, recommendations, and success stories; we basically use the same methods we were using when Mann and Barnard were teaching the nation the ideas of a universal education. This inability to train teachers becomes evident when teachers are assigned to work with the disadvantaged. The limited success of teachers seems linked more with personality than with training (Ornstein, 1969c). Research on teacher behavior should be conducted in order to get away from the "technique," "story," "hit-or-miss" approach.

24. Criticism directed against teacher behavior and teaching, especially teachers who work with the disadvantaged, should cease, since no particluar pattern of behavior can be advocated.

25. There is a need to improve rapport and communication between professors and teachers; moreover, invesitgators should provide clear and comprehendible feedback of their findings to teachers and school officials.

insert next page →

26. ~~Worthwhile and useful teacher-behavior should (1) enhance the feasiblity of determining merit pay and (2) enhance the status of the teaching profession.~~ In

26a In view of the growing demand for student power, especially at the college level, the value of (or lack of value) student evaluation of teaching performance should be judiciously considered. Do students have the right to evaluate faculty members? Are faculty members reasonably receptive to student evaluation? Do students' evaluations have a positive and continuing effect upon improving teaching (Academe, 1970)? What use will be made of the evaluations. These are serious questions, but what is more germane to our discussion is the reliability and validity of such evaluations? Several factors ~~also~~ need to be considered, too. Space limitations permit the investigator to mention only a few: (1) ~~the~~ type of course - required or elective; (2) size of student enrollment - a ten student seminar or a one hundred student lecture; (3) professor's distribution of grades; (4) degree of student unrest or dissent;(5) differences in students' and professor's social and political philosophy; (6) distinction between teacher behavior and course objectives; (7) professors notoriety (a "halo effect" might affect the students' rating; (8) differences in time (9 AM or 7 PM), days (Monday or Saturday), or semesters (Fall or Summer; (9) professor's teaching load; (10) professor's extent of secretarial or student assistance (ability to provide an abudance of mimeographed materials and quickly grade and return tests).

Until rating scales of teacher behavior are considered more reliable and valid, they should not be used for any purpose except for personal feedback. A "good" teacher or professor has little to worry about but can learn from his students. A "poor" instructor needs to know what students feel.

26b Another trend that seems to be fermenting within the teaching profession, and especially directed toward inner-city schools, is the question of teacher "accountability." The question of who has the right and expertise to rate teachers may soon challenge the teachers' probity, perhaps pitch teachers into conflict with students, parents, and/or supervisors. Similar questions, variables, and conclusions, as previously mentioned with regard to professors, also emerge.

26c However, Worthwhile and useful teacher-behavior ratings should (1) enhance the feasibility of determining merit pay and (2) enhance the status of the teaching profession.

27a Whether teachers are anti-research may no longer be a problem; in response to widespread criticism, they seem to be increasingly anti-researchers. This problem seems to be compounded by growing tension within the inner-city schools, whereby schools may be confronted with too many student-teacher-parent-administrator problems to risk having an outsider or researcher conduct a ~~research~~ study project or perceive the school's plight - perhaps report it in the study or write a commerical success. Similarly, in context with the black power movement, northern schools that are controlled, in part, by black communities, or which have a large, militant black staff, will probably cease welcoming white investigators.

27b In theory, almost any study conducted by a white investigator about the black community, school, or child can be ~~xxx~~construed as a potential Moynihan Report or Jensen exploration. Racial minorities and/or the educational establishment need not have to accept such findings, but they should acknowledge them. In view of the black-white conflict, it is problematical whether educators should or could pursue or accept research findings that negatively depicts any racial or ethnic minority group. On the other hand, hypersensitive egalitarianism may be the worst opponent of frank discussion and may impede racial equality. As Moynihan (1968) suggests, until more social scientists are drawn from minority groups and avialable to take part in research projects, the research establishment (including doctoral students who wish to study the disadvantaged) ~~xxxxxxxxxx~~ will probably be inhibited to explore such areas.

Insert

26.

In view of the growing demand for student power, especially at the college level, the value or lack of value in student evaluation of teaching performance should be judiciously considered. Do students have the right to evaluate faculty members? Are faculty members reasonably receptive to student evaluation? Do student students evaluations have a positive and continuing effect upon improving teaching? What use is to be made of the evaluations? These are serious questions, but they are not germane to this paper. What is linked to our discussion, however, is: how reliable and valid are such evaluations. Granted, that rating scales of teacher behavior, administered to students, may be suspect, the author feels that a "good" teacher or professor has little to worry about but can learn from his students. If a college or university feels that student evaluation is worthwhile, it behooves the institution to do it well.

* Academe, 1970, 4.

REFERENCES

Academe, February 1970, N.P.

Ackerman, W. J. Teacher competence and pupil change. Harvard Educational Review, 1954, 24, 273-289.

Allport, G. W. Attitudes. In. C. Murchison (Ed.), A handbook of social psychology. Worcester, Massachusetts: Clark University, 1935. Pp. 798-844.

American Educational Research Association, Committee on the Criteria of Teacher Effectiveness. Report of the committee on the crietia of teacher effectiveness. Review of Educational Research, 1952, 22, 238-263.

Anastasi, A. Psychology, psychologists, and psychological testing. American Psychologist, 1967, 22, 297-306.

Anderson, H. M. A study of certain criteria of teaching effectiveness. Journal of Experimental Education, 1954, 23, 41-71.

Atkin, J. M. Research styles in science education. Journal of Research in Science Teaching, 1967 - 1968, 5, 338-345.

Barr, A. S. Summary and comments. Journal of Experimental Education, 1946, 15, 99-100.

Barr, A. S. Teacher competencies. In W. S. Monroe (Ed.), Encyclopedia of Educational Research. (rev. ed.) New York: Macmillan, 1950. Pp. 1446-1454.

Barr, A. S., Eustice, D. E., & Noe, E. J. The measurement and prediction of teacher efficiency. Review of Educational Research, 1955, 25, 261-269.

Beck, W. R. Pupils' perceptions of teacher merit: A factor analysis of five postulated dimensions. Journal of Educational Research, 1967, 61, 127-128.

Bellack, A. A. & Others. The language of the classroom. New York: Teachers College, Columbia University, 1966.

Berdie, R. F. The uses of evaluation in guidance. In R. W. Tyler (Ed.), Educational evaluation: New roles, new means. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1969. Pp. 51-80.

Berelson, B., & Steiner, G. A. Human behavior: An inventory of scientific findings. New York: Harcourt, 1964.

Bettelheim, B. The decision to fail. School Review, 1961, 69, 389-412.

Biddle, B. J. The integration of teacher effectiveness reasearch. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 1-40.

Biddle, B. J. Methods and concepts in classroom research. Review of Educational Research, 1967, 37, 337-357.

Biddle, B. J., & Ellena, W. J. Preface. In B. J. Biddle & W. J. Ellena (Eds.),

Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. v-vii.

Bloom, B. S. Stability and change in human characteristics. New York: Wiley, 1964.

Bloom, B. S. Some theoretical issues relating to educational evaluation. In R. W. Tyler (Ed.), Educational evaluation: New roles, new means. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1969. Pp. 26-50.

Bloom, B. S. (Ed.), Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. Taxonomy of educational objectives, the classification of educational goals, handbook I: Cognitive domain. New York: McKay, 1956.

Borg, W. R. Personality and interest measures as related to criteria of instructor effectiveness. Journal of Educational Research, 1957, 50, 701-710.

Broudy, H. S. Can we define good teaching. Record, 1969, 70, 583-592.

Brown, B. B. Observer-judge ratings of teacher competence. Childhood Education, 1967, 44, 205-207.

Bryan, R. C. Eighty-six teachers try evaluating student reactions to themselves. Educational Administration and Supervision, 1941, 27, 513-526.

Budd, W. C., & Blakely, L. S. Response bias in the Minnesota Teacher Attitude Inventory. Journal of Educational Research, 1958, 51, 707-709.

Buros, O. K. (Ed.), Sixth mental measurements yearbook. Highland Park, New Jersey: Gryphon, 1965.

Callis, R. Change in teacher-pupil attitudes related to training and experience. Educational and Psychological Measurements, 1950, 10, 718-727.

Carras, W. G. Coversation with this author. New York University, Department of Social Studies Education, December 17, 1969.

Carroll, J. B. Words, meanings and concepts. Harvard Educational Review, 1964, 34, 178-202.

Christensen, C. M. Relationships between pupil achievement, pupil affect-need, teacher warmth, and teacher permissiveness. Journal of Educational Psychology, 1960, 51, 169-174.

Clark, K. B. Dark Ghetto. New York: Harper, 1965.

Cliff, N. Adjective check list responses and individual differences in perceived meaning. Educational and Psychological Measurements, 1968, 28, 1063-1078.

Cogan, M. L. The behavior of teachers and the productive behavior of their pupils: II trait analysis. Journal of Experimental Education, 1968, 27, 89-105.

Coleman, W. Susceptibility of the Minnesota Teacher Attitude Inventory to "faking" with experienced teachers,. Educational Administration and Supervision, 1954, 40, 234-237.

Cook, D. L. The hawthorne effect in educational research. Phi Delta Kappan, 1962, 44, 116-122.

Crawford, P. L., & Bradshaw, H. L. Perception of characteristics of effective university teachers: A scaling analysis. Educational and Psychological Measurement, 1968, 28, 1079-1085.

Cronbach, L. J. Essentials of psychological testing (2nd ed.) New York: Harper, 1959.

Cronbach, L. J. Evaluation for course improvement. Teachers College Record, 1963, 64, 672-683.

Cronbach, L. J. The role of the university in improving education. Phi Delta Kappan, 1966, 47, 539-545.

Cronbach, L. J. How can instruction be adapted to individual differences? In R. M. Gagne (Ed.), Learning and individual differences. Columbus, Ohio: Merrill, 1967. Pp. 23-39.

Crowne, D. P., & Marlowe, D. A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 1960, 24, 349-354.

Davis, A. Social-class influences upon learning. Cambridge: Harvard University, 1948.

Deutsch, M., Fishman, J. A., Kogan, L, North, R., & Whiteman, M. Guidelines for testing minority group children. Journal of Social Issues, 1964, 22, 127-145.

Drawhorne, C. L. Relationship between pupil and student-teacher interaction and pupil ratings of teacher effectiveness. Educational Administration and Supervision, 1954, 40, 283-296.

Dumas, W. W. Strengths and weaknesses of student teachers in English. Journal of Experimental Education, 1966, 35, 19-27.

Ebel, R. L. The relation of testing programs to educational goals. In W. G. Findley (Ed.), The impact and improvement of school testing programs. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1963. Pp. 28-44.

Ebel, R. L. Some limitations of basic research in education. Phi Delta Kappan, 1967, __, 81-84.

Eddy, E. M. Walk the white line: A profile of urban education. Garden City, New York: Doubleday, 1967.

Edwards, A. L., & Diers, C. J. Social desirability and the factorial interpretation of the MMPI. Educational and Psychological Measurement, 1962, 22, 501-509.

Eisner, E. W. Qualitative intelligence and the act of teaching. Elementary School Journal, 1963, 63, 299-307.

Erikson, E. H. Childhood and society. New York: Norton, 1950.

Flanders, N. A. Diagnosing and utilzing social structures in classroom learning. In N. B. Henry (Ed.), The dynamics of instructional groups. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1960. Pp. 187-217.

Flanders, N. A. Some relationships among teacher influence, pupil attitudes, and achivement. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 196-231.

Flanders, N. A. Teacher influence, pupil attitudes, and achivement. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1965. Research Monograph No. 12.

Flanders, N. A. Interaction analysis and inservice training. Journal of Experimental Education, 1968, 37, 126-133.

Foa, U. G. New developments in facet design and analysis. Psychological Review, 1965, 72, 267-274.

French, R. L. Research as a basis for creative teaching. Educational Horizons, 1961, 39, 28-34.

Gage, N. L. Paradigms for research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 94-141.

Gage, N. L. An analytical approach to research on instructional methods. Phi Delta Kappan, 1968, 49, 601-606.

Gage, N. L., Runkel, P. J., & Chaterjee, B. B. Equilibrium theory and behavior change: An experiment in feedback from pupils to teachers. Urbana: University of Illinois, 1960. Mimeographed.

Gage, N. L., & Unruh, W. R. Theoretical formulations for research on teaching. Review of Educational Research, 1967, 37, 358-370.

Gallagher, J. J. Productive thinking of gifted children. Urbana: University of Illinois, 1965. U. S. Department of Health, Education, and Welfare. Research Project No. 965.

Gallagher

Gallagher, J. J., & Aschner, M. J. A preliminary report on analysis of classroom interaction. Merrill-Palmer Quarterly, 1963, 9, 183-194.

Galloway, C. M. An exploratory study of observational procedures for determing teacher nonverbal communication. Unpublished doctoral dissertation. University of Florida, 1962.

Galloway, C. M. Nonverbal communication in teaching. Educational Leadership. 1966, 24, 55-63.

Gephart, W. J., & Antonoplos, D. P. The effects of expectancy and other research-biasing factors. Phi Delta Kappan, 1969, 50, 579-583.

Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 506-582.

Getzels, J. W., & Thelen, H. A. The classroom group as a unique social system. In N. B. Henry (Ed.), The dynamics of instructional groups. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1960. Pp. 53-82.

Gibb, J. R. Sociopsychological processes of group instruction. In N. B. Henry (Ed.), The dynamics of instructional groups. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1960. Pp. 115-135.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 1963, 18, 519-521.

Glaser, R. Toward a behavioral science for instructional design. In R. Glaser (Ed.), Teaching machines and programmed learning, II: Data and directions. Washington, D. C.: NEA, 1965. Pp. 771-809.

Goheen, R. F. The teacher in the university. American Scientist, 1966, 54, 221-225.

Gordon, L. V. Gordon personal profile manual. (Rev. ed.) New York: Harcourt, 1963.

Guba, E. G., & Getzels, J. W. Personality and teacher effectiveness: A problem in theoretical research. Journal of Educational Psychology, 1955, 46, 330-334.

Guilford, J. P. Psychometric methods. (2nd ed.) New York: McGraw-Hill, 1954.

Guilford, J. P. Intelligence: 1965 model. American Psychologist, 1966, 21, 20-26.

Gump, P. V. Environmental guigance of the classroom behavioral system. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 165-195.

Gump, P. V. Setting variables and the prediction of teacher behavior. Paper presented to the American Educational Research Association, New York, February, 1967.

Hargadon, B. K. Transference: A student-teacher interaction. School Review, 1966, 74, 446-452.

Harvey, O. J., Prather, M., White, B. J.,&Hoffmeister, J. K. Teachers' beliefs, classroom atmosphere and student behavior. American Educational Research Journal, 1968, 5, 151-166.

Havighurst, R. J. Developmental tasks and education. New York: Longmans, 1950.

Havighurst, R. J. Requirements for a valid "new criticism." Phi Delta Kappan, 1968, 50, 20-26.

Hawkins, E. E., & Stoops, E. Objective and subjective identification of outstanding elementary teachers. Journal of Educational Research, 1966, 59, 344-346.

Hearn, A. The relationship between specific traits and general teaching competence. Educational Administration and Supervision, 1953, 39, 500-503.

Hedges, W. D., & Macdougall, M. A. Recording student perceptions of teacher behavior by mean of a student response monitor. Journal of Educational Research, 1964, 58, 163-166.

Henry, J. Docility, or giving the teacher what she wants. Journal of Social Issues, 1955, 2, 33-41.

Himmelfarb, S. The stability of attitude item scale values. Journal of Social Psychology, 1969, 77, 107-111.

Hoffman, B. The tyranny of testing. New York, Crowell-Collier, 1964.

Hollingworth, H. L. Judging human character. New York: Appleton-Century-Crofts, 1922.

Holt, J. How children fail. New York: Pitman, 1964.

Hughes, M. M. What is teaching? One viewpoint. Educational Leadership, 1962, 19, 251-259.

Hughes, M. M., & Associates. Devel pment of the means for the assessment of the quality of teaching in elementary schools. In M. D. Usdan & F. Bertolaet (Eds.), Development of school-university-programs for the pre-service education of teachers for the disadvantaged through teacher education centers. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1965. Pp. 159-167. Research Project No. F-068.

Hunt, H. C. The ideal teacher. Journal of Education, 1942, 125, 37-38.

Hunt, J. M. Intelligence and experience. New York: Ronald, 1961.

Hunt, J. M. The psychological basis for using pre-school enrichment as an antidote for cultural deprivation. Merrill-Palmer Quarterly, 1964, 10, 209-248.

Hyman, R. T. Introduction. Teaching: Vantage points for study. Philadelphia: Lippincott, 1968. Pp. 1-12.

Jackson, J. M. Structural characteristics of norms. In N. B. Henry (Ed.), The dynamics of instructional groups. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1960. Pp. 136-163.

Jackson, P. W. The way teaching is. In O. Sand & L. J. Bishop (Eds.), The way teaching is. Association for Supervision and Curriculum Development. Washington, D. C.: ASCD, 1966. Pp. 7-27

Jackson, P. W. Life in the classroom. New York: Holt, 1968.

James, W. Talks to teachers on psychology (3rd ed.) New York: Norton, 1958. (First published in 1899.)

Jenkins, D. H. Characteristics and functions of leadership in instructional groups. In N. B. Henry (Ed.), The dynamics of instructional groups. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1960. Pp. 164-184.

Jensen, G. E. The social structure of the classroom group: An observational framework. Journal of Educational Psychology, 1955, 46, 362-374.

Jersild, A. T. Characteristics of teachers who are "liked best" and "disliked most." Journal of Experimental Education, 1940, 9, 139-151.

Johnson, D. W. Influences on teachers' acceptance of change. Elementary School Journal, 1969, 70, 142-153.

Johnson, D. G., Jr. An experimental technique for the prediction of teacher effectiveness. Journal of Educational Research, 1957, 50, 679-689.

Justiz, T. B. A reliable measure of teacher effectiveness. Educational Leadership, 1969, 3, 49-55.

Keislar, E. R., & McNeil, J. D. The use of pupil accomplices to investigate teacher behavior. Journal of Experimental Education, 1959, 27, 237-240.

Kerlinger, F. N., The factor structure and content of perception of desirable characteristics of teachers. Educational and Psychological Measurement, 1967, 27, 643-656.

Kerlinger, F. N., & Kaya, E. The predictive validity of scales constructed to measure attitudes toward education. Educational and Psychological Measurements, 1959, 1959, 305-317.

Klein, J. T. Presuppositions of teaching. Educational Theory, 1969, 19, 299-307.

Kohlberg, L. Early education: A cognitive-developmental view. Child Development, 1968, 39, 1013-1062.

Lamke, T. A. Introduction. Review of Educational Research, 1955, 25, 192.

Leacock, E. B. Teaching and learning in city schools. New York: Basic Books, 1969.

Lehmann, C. F. Analyzing and managing the physical setting of the classroom group. In N. B. Henry (Ed.), The dynamics of instructional groups. National Society for the Study of Education, Part II. Chicago: University of Chicago, 1960. Pp. 253-267.

Maccoby, E., & Maccoby, N. The interview: A tool of science. In G. Lindzey (Ed.), Handbook pf social psychology, Vol. II. Cambridge: Addison-Wesley, 1954. Pp. 449-487.

Macdonald, J. B., & Zaret, E. A study of openness in classroom interaction. In R. T. Hyman (Ed.), Teaching: Vantage points for study. Philadelphia: Lippincott, 1968. Pp. 197-208.

Maisa, B. B. Evaluating educational outcomes by means of formal behavioral science instruments. In M. D. Usdan & F. Bertolaet (Eds.), Development of school university-programs for the pre-service education of teachers for the disadvantaged through teacher education centers. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1965. Pp. 214-251.

Manning, W. H. The measurement of intellectual capacity and performance. Journal of Negro Education, 1968, 37, 258-267.

Masling, J., & Stern, G. Effect of the observer in the classroom. Journal of Educational Psychology, 1969, 60, 351-354.

McCallon, E. L. Interpersonal perception characteristics of teachers. Journal of Experimental Education, 1966, 34, 97-100.

McCallon, E., & Dumas, W. A factorial comparison of three teacher interpersonal perception measures. Journal of Educational Research, 1967, 61, 19-21.

McNemar, Q. On growth measurements. Educational and Psychological Measurements, 1958, 18, 47-55.

Medley, D. M. Teacher personality and teacher-pupil rapport. Journal of Teacher Education, 1961, 12, 152-156.

Medley, D. M., & Mitzel, H. E. A technique for measuring classroom behavior. Journal of Educational Psychology, 1958, 49, 86-92.

Medley, D. M., & Mitzel, H. E. Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 247-328.

Meux, M., & Smith, B. O. Logical dimensions of teaching behavior. In B. J. Biddle & W. L. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 127-164.

Michaelis, J. U. The prediction of success in student teaching from personality and attitude inventories. University of California Public School Education, 1954, 11, 415-481.

Morris, S. Teaching oractice: Objectives and conflicts. Educational Review, 1969, 21, 120-129.

Moynihan, D.P. Sources of resistance to the Coleman Report. Harvard Educational Review, 1968, 38, 23-36.

Oelke, M. C. A study of student teachers' attitude toward children. Journal of Educational Research, 1955, 48, 711-714.

Orleans, J. S., Clarke, D., Ostreicher, L, & Standlee, L. Some preliminary thoughts on the criteria of teacher effectiveness. Journal of Educational Research, 1952, 45, 641-648.

Orne, M. T. On the social psychology of the psychological experiment with particular reference to demand characteristics and their implications. American Psychologist, 1962, 17, 776-783.

Ornstein, A. C. Preparing and recruiting teachers for slum schools. Journal of Secondary Education, 1967, 42, 368-374.

Ornstein, A. C. Anxieties and forces which mitigate against ghetto school teachers. Journal of Secondary Education, 1968, 43, 243-254.

Ornstein, A. C. Discipline practices for teaching the disadvantaged. In A. C. Ornstein & P. D. Vairo (Eds.), How to teach disadvantaged youth. New York: McKay, 1969. Pp. 163-193. (a)

Ornstein, A. C. Letter to Stanley Holwitz, Editor, D. C. Heath & Co., December 28, 1969.(b)

Ornstein, A. C. Theory practices for teaching disadvantaged youth. Journal of Negro Education, 1969, 38, 82-85. (c)

Ornstein, A. C. Letter to Hulda Grobman, Professor of Secondary Education, New York University, January 28, 1970. (a)

Ornstein, A. C. Review of E. B. Leacock. Teaching and learning in city schools. Harvared Educational Review, 1970, 40, in print. (b)

Ornstein, A. C. Urban teachers and schools: Critics, conflict, and challenge. Educational Forum, 1971, 35, in print.

Packer, C. K., & Packer, T. Cybernetics, information theory and the educative process. Teachers College Record, 1959, 61, 134-142.

Paraskevopoulous, I. How students rate their teachers. Journal of Educational Research, 1968, 62, 25-29.

Perkins, H. V. A procedure for assessing the classroom behavior of students and teachers. American Educational Research Journal, 1964, 1, 249-260.

Phillips, J. A., Jr., Teacher typologies. High School Journal, 1967, 51, 26-31.

Rambo, L. M. Effects of experimental bias on attitudes toward controversial issues. Journal of Negro Education, 1969, 38, 384-394.

Rees, R. D. Dimensions of students' points of view in rating college teachers. Journal of Educational Psychology, 1969, 60, 476-482.

Remmers, H. H. Rating methods in research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 329-378.

Rivlin, H. N. New teachers for new immigrants. Teachers College Record, 1965, 66, 707-718.

Rogers, C. R. Significant learning: In therapy and in education. Educational Leadership, 1959, 16, 232-242.

Rosencranz, H. A., & Biddle, B. J. The role approach to teacher competence. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 232-263.

* Peterson, W. A. Age, teacher's role, and the institutional setting. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 232-263.

Rosenthal, R., & Jacobson, L. Pymaglion in the classroom. New York: Holt, 1968.

Ryans, D. G. A study of criterion data (A factor analysis of teacher behavior in the elementary school). Educational and Psychological Measurement, 1952, 12, 333-344.

Ryans, D. G. Characteristics of Teachers. Washington, D. C.: American Council on Education, 1960.

Ryans, D. G. Assessment of teacher behavior and instruction. Review of Educational Research, 1963, 33, 415-441. (a)

Ryans, D. G. Teacher behavior theory and research: Implications for teacher education. Journal of Teacher Education, 1963, 14, 274-293. (b)

Ryans, D. G. Research on teacher behavior in the context of the teacher characteristics study. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 67-101.

Scandura , J. M. Teaching-technology or theory. American Educational Research Journal, 1966, 3, 139-146.

Scheviakov, G. V., & Redl, F. Discipline for today's children and youth. (2nd ed.) Washington, D. C.: ASCD, 1956.

Siegel, L., & Siegel, L. C. The instructional gestalt. In L. Siegel (Ed.), Instruction: Some contemporary viewpoints. San Francisco: Chandler, 1967. Pp. 261-290.

Skinner, B. F. Reflections on a decade of teaching machines: In R. Glaser (Ed.), Teaching machines and programmed learning, II: Data and directions. Washington, D. C.: NEA, 1965. Pp. 5-20.

Smith, B. M. Conflicting values affecting behavioral research with children. Children, 1967, 11, 53-58.

Smith, B. O. Recent research on teaching: An interpretation. High School Journal, 1967, 51, 63-74.

Smith, B. O., and Others. A tentative report on the strategies of teaching. Urbana: University of Illinois, 1964. U. S. Department of Health, Education, and Welfare. Research Project No. 1640.

Smith, L. M., & Geoffrey, W. Toward a model of teacher decision-making in an urban classroom. St Louis: Washington University, 1965. U. S. Department of Health, Education, and Welfare. Final Report, Project No. S-048.

Smith, T. E. The image of high-school teachers: Self and others, real and ideal. Journal of Educational Research, 1965, 59, 99-104.

Sontag, M. Attitudes toward education and perception of teacher behaviors. American Educational Research Journal, 1968, 5, 385-402.

Sorenson, A. G. A note on the "fakeability" of the MTAI. Journal of Applied Psychology, 1956, 40, 192-198.

Sorenson, A. G., Husek, T. R., Yu, C. Divergent concepts of teacher role: An approach to the measurement of teacher effectiveness. Journal of Educational Psychology, 1963, 54, 287-294.

Sprinthall, N. A, Whiteley, J. M., & Mosher, R. L. Study of teacher effectiveness. Journal of Teacher Education, 1966, 17, 93-1-6.

Stake, R. E. The contenance of educational evaluation. Teachers College Record, 1967, 68, 523-540.

Start, K. B. The relation of teaching ability to measures of personality. British Journal of Educational Psychology, 1966, 36, 158-165.

Start, K. B. Rater-ratee personality in the assessment of teaching ability. British Journal of Educational Psychology, 1968, 38, 14-20.

Stavsky, W. H. Using the insights of psychtherapy in teaching. Elementary School Journal, 1957, 58, 28-35.

Stern, G. G. Measuring nongognitive variables in research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 398-447.

Stolurow, L. M. Model the master teacher or master the teaching model. In J. D. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand McNally, 1965. Pp. 223-247.

Strong, E. K., Jr. Vocational interests of men and women. Stanford, California: Stanford University, 1943.

Symonds, P. M. Characteristics of the effective teacher based upon pupil evaluations. Journal of Experimental Education, 1955, 23, 289-310.

Symonds, P. M., & Dudek, S. Use of the Rorschach in the diagnosis of teacher effectiveness. Journal of Projective Techniques, 1956, 20, 227-234.

Tanner, L. N. Teacher behavior and the destructive critics. School and Society, 1969, 97, 336-337, 370.

Taylor, J. B. Rating scales as measures of clinical judgment: A method for increasing scale reliability and sensitivity. Educational and psychological measurement, 1968, 28, 747-766.

Torrance, E. P. Teacher attitude and pupil perception. Journal of Teacher Education, 1960, 11, 97-102.

Travers, R. M. W. Towards taking the fun out out of building a theory of instruction. Teachers College Record, 1966, 68, 49-60.

Tschechtelin, M. A. As teacher sees teacher: A study in personality. Journal of Social Psychology, 1953, 38, 121-125.

Turner, R. L. Teaching as problem-solving behavior: A strategy. In B. J. Biddle & W. J. Ellena. Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 102-126.

Veldman, D. J., & Peck, R. F. Student teacher characteristics from the pupils' viewpoint. Journal of Educational Psychology, 1963, 54, 346-355.

Walberg, H. J. The development of teacher personality multivariate theory and analysis. School Review, 1967, 75, 187-196.

Walberg, H. J. Social environment as a mediator of classroom learning. Journal of Educational Psychology, 1969, 60, 443-448.

Wall, W. D. The future of educational research. Educational Research, 1968, 10, 163-169.

Warren, P. B. A study of lower class and middle-upper class students' perceptions of the behavior traits of the effective teacher. Unpublished doctoral dissertation. New York University, 1968.

Wehling, L. J., & Charters, W. W., Jr. Dimensions of teacher beliefs about the teaching process. American Educational Research Journal, 1969, 6, 7-30.

White, W. F., & Anderson, H. E. Jr. A study of scaled dimensions of teacher behavior as perceived by students. Journal of Psychology, 1967, 65, (second half) 223-232.

Wilkerson, D. A. The relation of research in psychology, sociology, and education to the instruction of disadvantaged pupils. In S. P. Koenigsberg (Ed.), Improving teacher education for disadvantaged youth: What university professors can learn from classroom teachers. New York: Yeshiva University, Conference Proceedings, May 15-17, 1966. Pp. 67-79. The author of the article is anonymous. However, this author telephoned Dr. Koenigsberg on April 24, 1969, and she indicated that Dr. Wilkerson wrote the article.

Williard, R. A. Discrepancies in learning experiences reported in classrooms. Educational Administration and Supervision, 1957, 43, 339-348.

Withall, J. The development of the climate index. Journal of Educational Research, 1951, 45, 93-100.

Yee, A. H. Source and direction of casual influence in teacher-pupil relationships. Journal of Educational Psychology, 1968, 59, 275-282.

Yourman, J. The case against group I. Q. testing in schools with culturally disadvantaged pupils. Phi Delta Kappan, 1964, 46, 108-110.

Zahorik, J. A. Classroom feedback of teachers. Journal of Educational Research, 1968, 62, 147-150.